2015 UPDATE #1 TO:

A METHODOLOGY TO DETERMINE ANNUAL EFFECTIVENESS OF EMISSION MITIGATION TECHNIQUES IN THE ODSVRA

By Mel Zeldin
Consultant to San Luis Obispo County APCD
and Karl A. Tupper
Air Quality Specialist, San Luis Obispo County APCD

Introduction

The original methodology and data were provided in December 2014 covering the years 2011 through 2014, with 2014 data available only for the first part of the year. Filter days were determined through a classification strategy that looks only at the 1000-1500 time period (hourly values) under stringent meteorological parameters that essentially hold meteorology constant such that PM10 concentrations measured at CDF are directly related to the emissions from the ODSVRA. The methodology was previously provided to the Technical Committee members. The specific criteria for defining a filter day are re-stated as follows:

During the six-hour period, 1000-1500:
1) All PM10, S1, and CDF wind speed and direction measurements must be valid;
2) The S1 vector average wind direction must be between 285 and 300 degrees for the six-hour period;
3) The S1 site must have all hourly wind speeds greater than or equal to 5 m/s;
4) The S1 site must have at least 3 of the six hourly wind speed greater than 10 m/s;
5) The S1 site must not have any hourly wind direction > 310 degrees;
6) The CDF site must not have any hourly wind direction < 285 degrees.

This update completes the data for stratified "filter days" for 2014, thus providing four-year "baseline" conditions from which future years with mitigations can be compared.

Also, this update includes validated data results through August 31 for 2015, which includes most of the primary windy period, and thus provides some insight into the 2015 mitigation effectiveness compared to the prior four-year baseline period.

Update of Statistical Estimates for the Baseline Period

With the completion of 2014 data, an additional four "filter days" were added to the baseline aggregate statistics. Also, three additional days (two in 2011, and one in 2013) were re-analyzed as meeting filter day criteria and now included in the baseline years. There are now 61 days aggregated over the four years baseline period. Because average wind speeds varied slightly year to year, annual data were normalized to wind speed to get an average concentration per m/s for each year. From these annual values, averages and standard deviations were determined, as shown below:

YEAR	# FILTER DAYS	CDF PM10	S1 WIND (m/S)	PM10/M PER S
2011	10	270	10.3	26.2
2012	16	357	11.7	30.5
2013	21	325	11.5	28.3
2014	14	317	10.7	29.6
	Avg	317.0	Avg	28.7
			Std Dev	1.9
			3 std dev	5.7
			Target 95% conf.	23.0
	Target level	253.9	% reduction needed	19.9%

Note that the normalized values for all four years are remarkably similar. 2011 has the fewest number of data "filter days" with an annual total of 10.

As a double check of the average normalized value as shown in the above table, all 366 data points for the 61 days over the four years were averaged and then normalized. The resulting value was 28.7 -- identical (as rounded) to the average of the annual four yearly values.

With the understanding that the standard deviation represents the inter-annual meteorological variability, to have confidence that an observed normalized annual value (ug/m3 PM10 per m/s wind) is statistically significant beyond the inter-annual variability, a value of 3 standard deviations was used as the confidence level necessary to determine that any observed reductions were due to mitigations, not meteorology. (The three sigma level was used because for a t-test with 3 degrees of freedom, the 95% confidence interval is 3 sigmas.) This results in a target normalized value of 23.0 ug/m3 per m/s, or 19.9% reduction from the mean value over the 4-year period. Applying that level of reduction to concentration means that the target level for the average 6-hour concentration of measured PM10 at CDF would need to be reduced to 254 ug/m3 for the average of all filter days during a year for the hours of 1000-1500. If that level is achieved with at least 10 "filter days," it can be determined that the reduction in PM10 has statistical confidence that it is related to the mitigations in place and not due to meteorological variability.

Evaluation of 2015 Through August

Data for 2015 are now available through August 2015.

The results for 2015 are shown below:

YEAR	# FILTER DAYS	CDF PM10 (ug/m3)	S1 WIND (m/s)	PM10/M PER S
2015	5	336	10.2	32.9

There are some very notable results for 2015. In order to better understand the annual distribution of "filter days", these are the occurrences by month for the 4-year baseline period:

March	3
April	17
May	23
June	12
July	2
August	0
September	3
October	1

(Other months are all zeroes)

As can be seen, the month with the most frequent "filter days" is May. However, May 2015 was a meteorologically anomalous month with a number of late winter/early spring storms parading through California. Its effects were pronounced, as there were zero "filter days" during May 2015. There were also zero filter days in June. The dates of the five "filter days" through the end of August are: March 31, April 1, April 4, April 26, and July 7 (all when wind fencing was either fully or mostly installed) . Because of the anomalous weather during May and June, the total number of filter days through August is only 5. In this analytical approach, it would be preferable to have at least 10 filter day events, similar to that which occurred in 2011. Because of the lower number of filter days in 2015, some caution is advised in the annual results. However, the results show normalized values of 32.9 ug/m3 per m/s wind speed at S1 which is1.8 standard deviations greater than the baseline normalized value. The increase in normalized values does run contrary to expected mitigation results, recognizing that these days occurred when most or all of the wind fencing was in place. One would have expected to see results on the lower side of the baseline normalized PM10 concentration as opposed to the observed levels 1.8 standard deviations above the baseline normal.

A plot of the annual wind speeds at S1 versus CDF PM10 values is shown in Figure 1.

An Alternative Statistical Approach

Because of the few number of filter days in 2015, another more rigorous statistical approach was taken to determine if any significance can be attached to the results for 2015.

This approach involves the analysis of variance (ANOVA) to determine between-group variance (annualized data) versus within-group variance in treating the baseline four-year data set as a pooled data set.

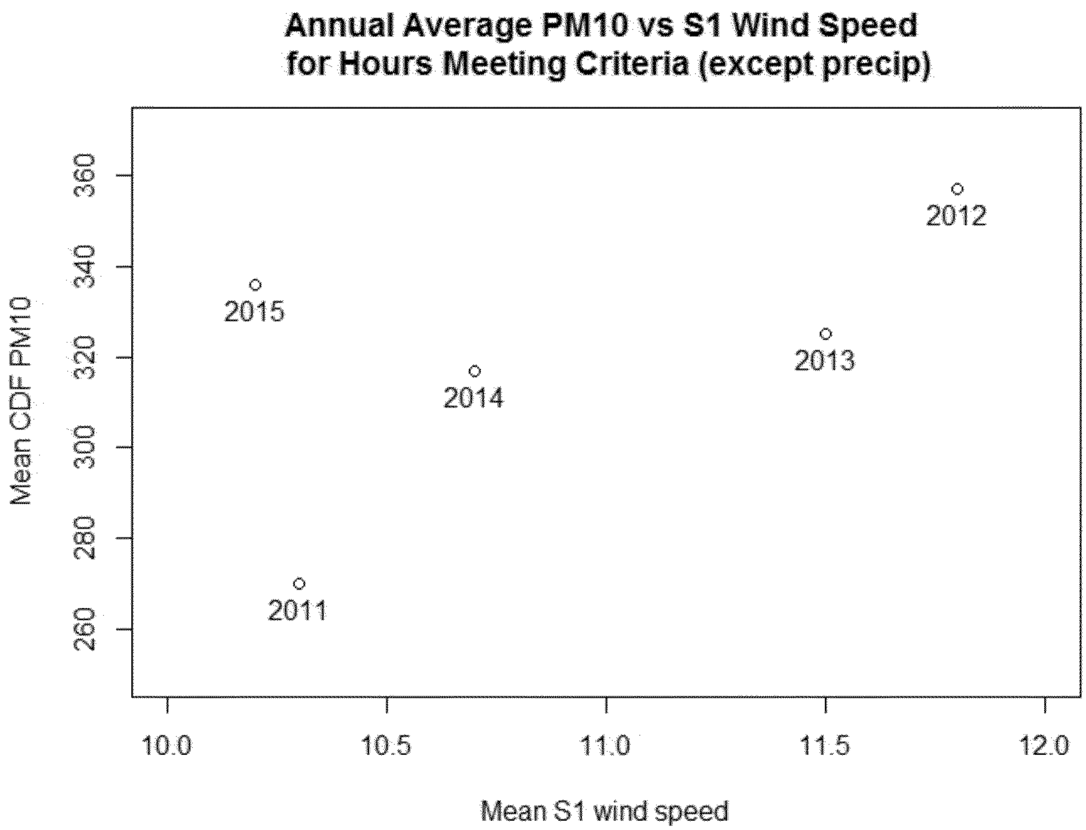

Figure 1. Annual CDF PM10 versus per m/s wind speed as recorded at the S1 tower. Note the fairly good linear fit for the baseline years (2011-2014) and the higher normalized values for the 5 days meeting filter day criteria in 2015.

Using the most recent dataset (396 hourly observations, with 10 filter days in 2011, 16 in 2012, 21 in 2013, 14 in 2014, and 5 this year), the PM10/WS for each hour (calling this new variable "ratio") was calculated. Data were grouped by year, and years 2011 to 2014 were analyzed by one way ANOVA to determine whether the years differed significantly from one another. The result is shown below:

```
Analysis of Variance Table

Response: ratio
           Df Sum Sq Mean Sq F value  Pr(>F)
year        3    569 189.811   2.142 0.09458 .
Residuals 362  32078  88.612
```

The p-value of 0.09458, which is not statistically significant, indicates that within-year variance is much greater than any between-year variance. In other words, the years (2011-2014) are not significantly different from each other, so it's appropriate to pool them.

Next these years were pooled together, and then the data for 2015 was compared to this baseline. The result is as follows:

```
Analysis of Variance Table

Response: ratio
             Df Sum Sq Mean Sq F value  Pr(>F)
year == 2015   1    353  353.33   3.884 0.04945 *
Residuals    394  35842   90.97
```

This is marginally statistically significant at the 95% confidence level, i.e. 2015 is different from the other years. How different is it? The output below indicates that the baseline PM10/WS value (labeled "(Intercept)") is 28.75 and 2015 is 3.57 units higher, and this difference is statistically significant, with a p-value of 0.0494.

```
Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       28.7453     0.4986  57.658   <2e-16 ***
year == 2015TRUE   3.5697     1.8113   1.971   0.0494 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.538 on 394 degrees of freedom
Multiple R-squared:  0.009762, Adjusted R-squared:  0.007248
F-statistic: 3.884 on 1 and 394 DF,  p-value: 0.04945
```

So the inference is that 2015 is significantly worse. Since this is a linear model, and difference of 3.57 is marginally significant at the 95% level, this implies that the ratio would have to get down 25.2 to be able declare a statistically significant improvement in air quality. (Whether this would be enough to comply with the rule is an entirely different question. Rather this indicates we can detect the effect, or lack thereof, of the wind fence mitigations on the CDF PM10). Note that this approach is equivalent to performing a T-test comparing 2015 to the pooled baseline years. This assumes equal variances, and an F-test suggests that the assumption of equal variances is valid (p-value =0.3883). ANOVA and T-tests both assume that the data is normally is distributed, and a histogram of the ratios looks normal (see below). Nonetheless, the analysis was repeated using the non-parametric Kruskal-Wallis rank sum test. This test also showed that the ratio for 2015 was significant higher than the baseline years (p-value = 0.031).

As an additional test, all five years, 2011 to 2015, were pooled, and then each year was pulled out and compared to the pooled data for the remaining four years. The results are given below:

```
> anova(lm(ratio~(year==2011),  noprecip))
Analysis of Variance Table

Response: ratio
             Df Sum Sq Mean Sq F value  Pr(>F)
```

```
year == 2011   1    345  344.61  3.7872 0.05236 .
Residuals    394  35851   90.99
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> anova(lm(ratio~(year==2012),  noprecip))
Analysis of Variance Table

Response: ratio
              Df Sum Sq Mean Sq F value  Pr(>F)
year == 2012   1    262 262.353  2.8766 0.09066 .
Residuals    394  35933  91.201
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> anova(lm(ratio~(year==2013),  noprecip))
Analysis of Variance Table

Response: ratio
              Df Sum Sq Mean Sq F value Pr(>F)
year == 2013   1    152  152.41   1.666 0.1975
Residuals    394  36043   91.48
> anova(lm(ratio~(year==2014),  noprecip))
Analysis of Variance Table

Response: ratio
              Df Sum Sq Mean Sq F value Pr(>F)
year == 2014   1      1   1.442  0.0157 0.9004
Residuals    394  36194  91.863
> anova(lm(ratio~(year==2015),  noprecip))
Analysis of Variance Table

Response: ratio
              Df Sum Sq Mean Sq F value  Pr(>F)
year == 2015   1    353  353.33   3.884 0.04945 *
Residuals    394  35842   90.97
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

As can be seen, only 2015 is statistically significant at the 95% confidence level as compared with the other four-year pooled data. Thus clearly, 2015 stands out uniquely as compared to the other years, with higher levels than observed on filter days in the baseline period. Why such an increase is observed in 2015 is unclear, but what is clear is that there is no indication that the wind fencing had any effect on reducing the PM10 levels observed at CDF during hours and days when the filter day criteria were met, and designed to capture nearly identical meteorological conditions conducive to the ODSVRA emissions impacting the CDF monitoring site.

PM10/WS ratio density and box plots for each year, 2011-2015 are shown in Figure 2.

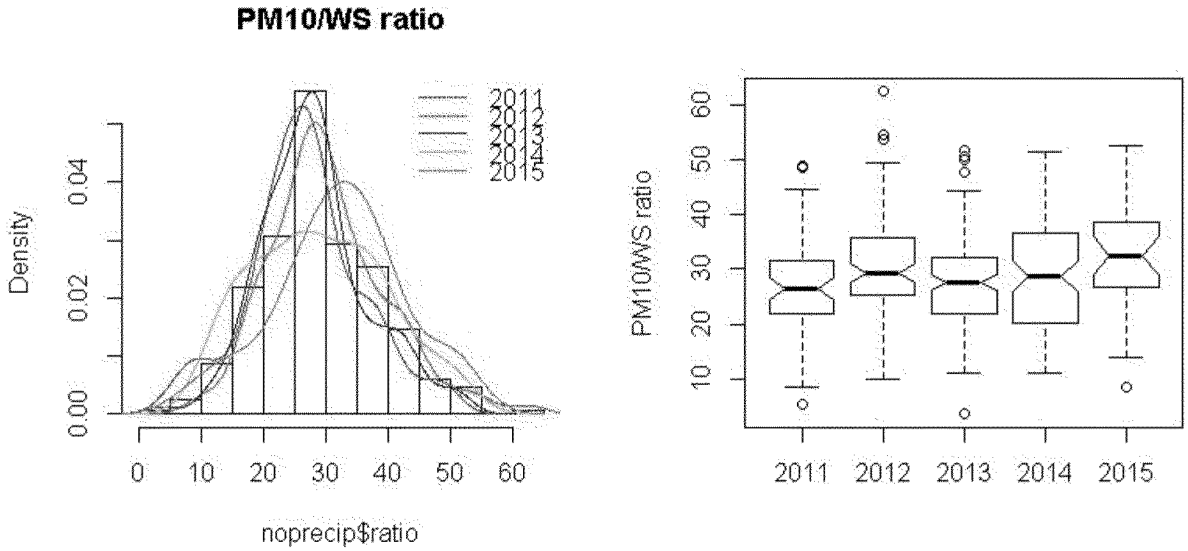

Figure 2. Density and box plots for PM10/WS ratio.